SEGMENTATION AND ESTIMATION OF THE HISTOLOGICAL COMPOSITION OF THE TUMOR MASS IN COMPUTED TOMOGRAPHIC IMAGES OF NEUROBLASTOMA

Fábio José Ayres[1,2], Marcelo Knörich Zuffo[2], Rangaraj Mandayam Rangayyan[1,2], Vicente Odone Filho[3], and Marcelo Valente[3]

[1]Department of Electrical and Computer Engineering, University of Calgary, Calgary, Canada
[2]Departamento de Engenharia de Sistemas Eletrônicos, Escola Politécnica da Universidade de São Paulo, São Paulo, Brazil
[3]Hospital das Clinicas, Instituto da Criança, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil

Abstract-**The problem that we investigate in the present paper is the improvement of the analysis of the primary tumor mass, in patients with advanced neuroblastoma, using X-ray computed tomography (CT) exams. To achieve this goal, we propose a methodology for the estimation of the histological content of the mass that comprises a technique for semi-automatic segmentation of the primary tumor mass in CT images of neuroblastoma and a statistical method to estimate, from segmented CT images, the histological composition of the primary tumor. The results of the method are compared with the results of histological analysis of surgically resected tumor mass.**

Keywords **- Image segmentation, computed tomography, feature extraction, neuroblastoma**

I. INTRODUCTION

Neuroblastoma is a malignant tumor of neural crest origin that may arise anywhere along the sympathetic ganglia or within the adrenal medulla. It is said to be an enigmatic and fascinating entity that represents, at the same time, one of the best examples of spontaneous regression for a malignant tumor and one of the tumors with the poorest cure potential for some groups [1]. Neuroblastoma is the most common extracranial solid malignant tumor in children; it is the third most common malignancy of childhood [2].

On computed tomography (CT) exams, abdominal neuroblastoma is seen as a mass of soft tissue, commonly suprarenal or paravertebral, irregularly shaped, lobulated, extending from the flank toward the midline, and lacking a capsule. The mass tends to be inhomogeneous due to tumor necrosis intermixed with viable tumor, and contains calcifications in 85% of patients. Calcifications are usually dense, amorphous, and mottled in appearance. Sometimes, neuroblastoma presents areas of central necrosis, shown as low-attenuation areas that are more apparent after contrast enhancement [2].

Despite the proven usefulness of imaging techniques in the detection, delineation, and staging of the primary tumor [2], there is a need for improvement in the usage of these techniques for a more accurate assessment of the local disease that could lead to better treatment planning and follow-up. Foglia et al. [3] argued that the primary tumor status in advanced neuroblastoma cannot be assessed definitively by diagnostic imaging, due to errors in sensitivity and specificity as high as 38%, when assessing tumor viability by imaging methods and comparing it to findings in delayed surgery. They also reported that CT exams could not differentiate viable tumor from fibrotic tissue or nonviable tumor destroyed by previous chemotherapy.

Computer-based image analysis and other techniques could improve radiological analysis of neuroblastoma by offering more sophisticated, accurate, and reproducible measures of information in the image data [4]. Nevertheless, few researchers have investigated the potential of computer-aided diagnosis of neuroblastoma using diagnostic imaging; related published works are limited to tumor volume measurement using manually segmented CT slices (planimetry) [5]. In particular, there is no method, with or without computer aid, that could assess with precision, using image-based exams, the amount of active disease, scar tissue, fibrosis, benign tumor, and other tissues in a given tumor mass [3].

II. METHODOLOGY

The problem we propose to investigate is quantitative analysis of the primary tumor mass, in patients with advanced neuroblastoma, using CT exams. To achieve this goal, we propose a methodology for the estimation of the histological content of the mass, that comprises a technique for semi-automatic segmentation of the tumor mass, a statistical parametric model for the histological composition of the tumor, and a method to estimate the parameters of the model. The segmentation algorithm implemented is based on the fuzzy connectivity concept as defined by Udupa and Samarasekera [6]. The statistical model employed is the Gaussian mixture model, and the algorithm for parameter estimation is the Expectation-maximization algorithm [7].

The complete methodology is shown in Figure 1. The path on the left shows the sequence of computer-based operations that lead to the estimation of the statistical model. The path on the right consists of obtaining the histological information regarding the tumor from delayed surgery and pathological analysis. Correlation of the information drawn from the two approaches should prove if the model is appropriate.

A. Segmentation

The problem of automatically or semi-automatically segmenting neuroblastoma using medical image processing techniques has not been studied, to the best of our knowledge.

This project has been supported by grants from the Kids' Cancer Care Foundation of Alberta, Calgary, Alberta, Canada; the Natural Sciences and Engineering Research Council of Canada; and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), São Paulo, Brazil.

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25OCT2001	N/A	-

Title and Subtitle Segmentation and Estimation of the Histological Composition of the Tumor Mass in computed Tomographic Images of Neuroblastoma	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Department of Electrical and Computer Engineering, University of Calgary, Calgary, Canada	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research Development & Standardization Group (UK) PSC 803 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from the 23rd Annual International conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.

Abstract

Subject Terms

Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU

Number of Pages
4

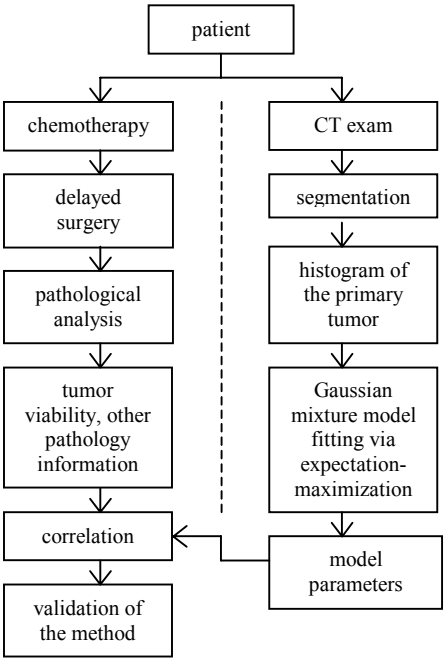

Fig. 1: Methodology for computer-based analysis of neuroblastoma.

The insight for the design of the techniques in the present work comes from other segmentation problems in the medical image processing area. The technique that we investigated for semi-automatic segmentation is based on the concept of *fuzzy connectivity* as developed in the work of Udupa and Samarasekera [6]. The fuzzy connectivity algorithm is a region growing procedure whose output, for each volume element (voxel) or picture element (pixel) in the image, is a value in [0,1] that represents the fuzzy degree of connectivity between the voxel and the seed voxel (or seed region) that is used to commence the procedure.

Let C represent a set of features of each voxel, which is comprised of the spatial location of the voxel and a few local properties (such as local texture, CT value, and gradient). Let us define a function $\eta : C \times C \rightarrow [0,1]$ (a fuzzy relation) that represents the fuzzy affinity between two voxels. The function η has the following properties: $\eta(a\ ,\ a) = 1$ (the relation is reflexive) and $\eta(a\ ,\ b) = \eta(b\ ,\ a)$ (the relation is symmetric). Qualitatively, the affinity between two voxels must be high if they are close in space and if their properties are similar, i.e., if they are close in the feature space.

A path of size n between two voxels a and b is a sequence $\langle c_1 = a\ ,\ c_2\ ,\ \ldots\ ,\ c_n = b\rangle$ of voxels $\{c_i\}$. The strength of the path is $\min[\eta(c_i\ ,\ c_{i+1})]$. The *connectedness* between a and b is the strength of the strongest path between them, among all possible paths. Finally, a fuzzy object can be defined given a seed voxel and the function η evaluating for each voxel the fuzzy connectedness between each voxel and the seed voxel (or the seed region). Udupa and Samarasekera [6] have shown how this calculation can be performed using dynamic programming.

The method described above was implemented in our work, with $\eta(a\ ,\ b) = 0$ if voxels a and b are not neighbors, and

$$\eta(a,b) = \exp\left[-\frac{1}{2\pi\sigma^2}\left(\frac{f(a)+f(b)}{2}-\mu\right)\right] \qquad (1)$$

if a and b are neighbors, where $f(a)$ is the CT value of the voxel a, and μ and σ are the mean and the standard deviation of the region. The values of μ and σ are estimated from the seed region.

In addition to the above, tumor regions were also segmented manually by an experienced radiologist (MV), on each CT slice. Manual segmentation represents the expert knowledge for comparison with the results of image processing.

B. Histogram Evaluation

In a CT image, each voxel contains an integer number which is the CT value, in Hounsfield units (HU), of the voxel. Therefore, the standard method of histogram evaluation was used with the manually segmented tumor mass.

However, given the fuzzy object computed in the semi-automatic segmentation procedure, we must evaluate the histogram in a different way. In this case, each voxel has a fuzzy membership value that quantifies the degree to which each voxel belongs to the tumor mass. To evaluate the histogram we proceed as follows: for each voxel, we add its membership value (a real number between zero and one) to the histogram position corresponding to the CT value of the voxel.

The two histograms obtained as above were compared. The histogram obtained from the result of manual segmentation is considered to be the true histogram of the tumor for comparison purposes.

C. Estimation of Histological Composition

Once the primary neuroblastoma mass is segmented and the histogram obtained, we proceed to the estimation of its histological composition. The tumor mass is inhomogeneous, due to intermixed necrosis and viable tumoral tissue, and sometimes presents central areas of necrosis, shown as low-attenuation regions inside the mass [2]. Therefore, we need to develop a global description of the mass that could lead to the estimation of the fractional volume corresponding to each tissue type, instead of attempting to separate the mass into distinct regions.

We assume that the CT value for a voxel that arises from a given type of tissue (benign mass, necrosis, malignant tumor, fibrotic tissue, etc.) inside the mass is a Gaussian random variable. Therefore, the whole tumor mass is modeled statistically as a mixture of Gaussian variables, known as the Gaussian mixture model [7].

Let x denote the CT attenuation value for a given voxel, and $\theta_i = (\mu_i\ ,\ \sigma_i)$ be the set of parameters that describes the

Gaussian probability density function of the CT values of the i^{th} type of tissue. The probability density function for x, given that x came from the i^{th} type of tissue, is $p_i(x \mid \theta_i) = N(\mu_i , \sigma_i)$, the one-dimensional normal distribution with mean μ_i and standard deviation σ_i.

Let M be the true number of different types of tissue in the given tumor mass, which we will assume to be known for the moment. Let α_i be the probability that a given voxel came from the i^{th} type of tissue in the tumor mass. By definition, the α_i values sum up to one. The value of α_i can be seen as the fraction of the tumor volume that is composed of the i^{th} type of tissue. Then, the probability density function for the entire mass is a mixture of Gaussians, specified by the parameter vector $\Theta = (\alpha_1 , \ldots , \alpha_M , \theta_1 , \ldots , \theta_M)$, and described by

$$p(x|\Theta) = \sum_{i=1}^{M} \alpha_i p_i (x|\theta_i). \quad (2)$$

Let N be the number of voxels in the tumor mass and let $\mathbf{x} = (x_1 , x_2 , \ldots , x_N)$ be a vector composed of the values of the voxels, which we will call the observed data. Given the observed data, we choose the best value of Θ to be the one that maximizes the likelihood of the data (the Maximum Likelihood principle), defined as

$$L(\Theta|\mathbf{x}) \equiv p(\mathbf{x}|\Theta) = \prod_{i=1}^{N} p(x_i|\Theta). \quad (3)$$

In order to maximize the likelihood, we use the Expectation-maximization (EM) algorithm [7]. The EM algorithm is an iterative procedure that starts with an initial guess Θ_g of the parameters and iteratively improves the estimate towards a local maximum of the likelihood. The generic EM algorithm is comprised of two steps: the expectation step (or E-step) and the maximization step (or M-step). In the E-step, one computes the parametric probability model given the current estimate of the parameter vector. In the M-step, one finds the parameter vector that maximizes the newly calculated model, which is then treated as the new best estimate of the parameters. The iterative procedure continues until some stopping condition is met, e.g., the difference $\log[L(\Theta_{i+1} \mid \mathbf{x})] - \log[L(\Theta_i \mid \mathbf{x})]$ or the modulus $|\Theta_{i+1} - \Theta_i|$ of the difference vector between successive iterations is smaller than a predefined value.

In order to investigate the effect of the number of estimated kernels (M) in our analysis, the EM algorithm was executed with several different values of M and the results were compared to one another and to the opinion of experts (VOF, MV).

III. Results

A three-year old female was diagnosed (in October, 2000) as having Stage IV neuroblastoma, with primary tumor located in the left adrenal gland, and with extensive metastatic disease in bone marrow and bones. After initial chemotherapy treatment there was complete disappearance of all metastases and an impressive diminishment of the primary tumor, which was then entirely removed surgically (in May, 2001). There was no detectable nodal disease. According to the pathology analysis, the tumor showed no more than 30% of viable tumor cells, the remaining 70% consisting of fibrosis and necrosis. The pre-surgical CT exam revealed a heterogeneous area, located in the left adrenal gland, composed of higher density areas (calcifications) and less dense areas (necrosis). After contrast injection, the radiologist (MV) observed a heterogeneous contrast uptake by the mass. No vascular structure infiltration was observed.

Figure 2 shows the tumor mass shrunk by chemotherapy, in a CT exam prior to surgical resection, with the tumor manually outlined by the radiologist (MV). The tumor mass was segmented using the fuzzy connectivity algorithm, and the result is shown in Figure 3. The normalized histogram of the tumor mass was then obtained from the manually and the automatically segmented exams. The results are shown in Figure 4, where the solid line corresponds to the normalized histogram obtained from the manually segmented dataset, and the dashed line shows the normalized histogram drawn from the automatically segmented tumor mass.

Figure 5a, 5b, and 5c show the result of the Gaussian mixture model estimation, using the EM algorithm, with the manually segmented tumor mass from the pre-surgery CT slices, and correspond to the estimation of two, three, and four Gaussian kernels, respectively. The estimated model parameters are shown in Table 1.

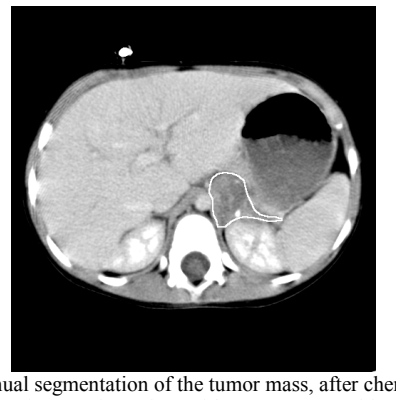

Fig. 2: Manual segmentation of the tumor mass, after chemotherapy treatment, in a CT image of a patient with Stage IV neuroblastoma (by MV).

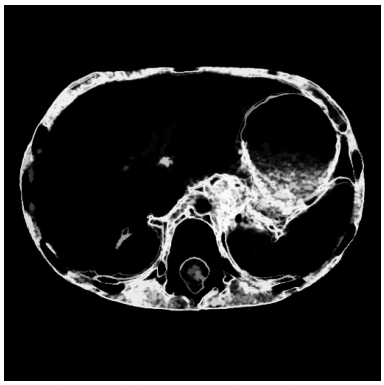

Fig 3: Fuzzy connectedness of the region obtained by region growing for the tumor illustrated in Figure 2.

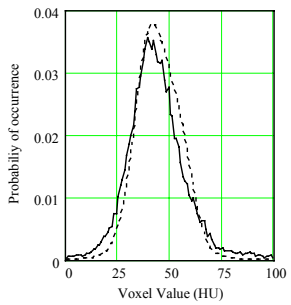

Fig. 4: Normalized histograms of the tumor mass, obtained from manual (solid line) and automatic (dashed line) segmentation of the pre-surgery CT images (33 slices).

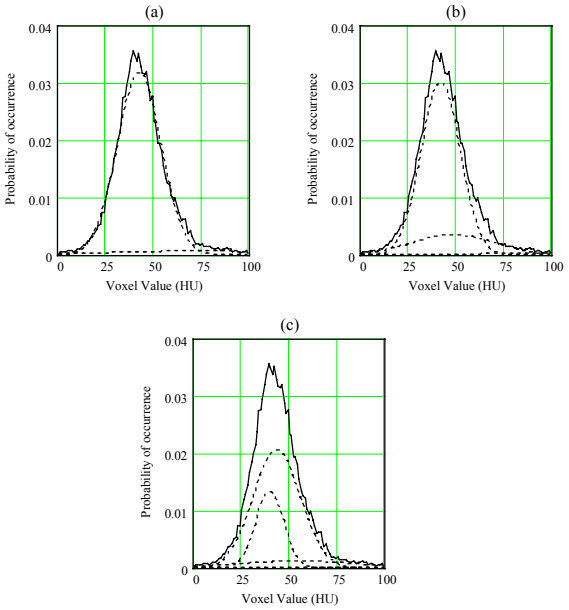

Fig. 5: Gaussian kernels estimated from the EM algorithm. Plots (a), (b), (c) show two, three and four estimated kernels (dashed lines), respectively, and the original histogram (solid) for comparison

IV. Discussion

Comparing Figures 2 and 3, it is seen that the fuzzy segmentation of the tumor mass does not correspond well with the manual segmentation: several areas outside the tumor mass are erroneously labeled as tumor. The fuzzy segmentation method needs to be improved or the result processed further to obtain better delineation of the tumor mass. While the normalized histograms presented in Figure 4 are similar, the corresponding parameters estimated for the Gaussian mixture model were significatively different (not shown).

Two medical experts analyzed the results of the Gaussian mixture model estimation, shown in Figure 5a, 5b, 5c, and Table 1: a radiologist (MV) and an oncologist (VOF). The results were compared to the pathology findings. From

TABLE 1
ESTIMATED PARAMETERS OF THE GAUSSIAN KERNELS

# of Gaussians	Weight	Mean (HU)	Standard Deviation (HU)
1	1.0	44.83	19.38
2	0.91	42.29	11.47
	0.09	71.82	46.61
3	0.77	41.69	10.17
	0.19	48.10	21.54
	0.04	93.60	54.77
4	0.63	43.31	12.35
	0.25	39.58	7.30
	0.11	54.09	31.13
	0.01	117.88	54.61

Table 1 it is seen that, when estimating the model with three Gaussian kernels, the kernel with the smaller mean value has a relative weight of 0.77 (77%). This result agrees with the post-surgical pathology analysis, which reported a minimum of 70% of fibrosis and necrosis, tissues that have a smaller value in HU than viable tumor.

Further work is in progress to improve the results of fuzzy segmentation, and to determine or select the number of Gaussian kernels to be used in the model. The methods will be tested with more cases as they become available. We will also explore the potential of the methods for analysis of CT scans during chemotherapy to estimate the change in the tumor in response to the treatment.

Acknowledgment

We thank Dr. Max Coppes of the Alberta Children's Hospital, Calgary, Alberta, Canada, for his advice on this project.

References

[1] F. Alexander, "Neuroblastoma", *Urologic Clinics of North America*, vol 27(3), pp. 383-392, 2000.

[2] S. J. Abramson, "Adrenal neoplasm in children", *Radiologic Clinics of North America*, vol 35(6), pp. 1415-1453, 1997.

[3] R. P. Foglia, E. W. Fonkalsrud, S. A. Feig, and T. J. Moss, "Accuracy of diagnostic imaging as determined by delayed operative intervention for advanced neuroblastoma", *Journal of Pediatric Surgery*, vol 24(7), pp. 708-711, 1989.

[4] I. N. Bankman, editor, *Handbook of Medical Image Analysis*, Academic Press, London, UK, 2000.

[5] J. M. Wheatley, N. S. Rosenfield, G. Heller, D. Feldstein, and M. P. LaQuaglia, "Validation of a technique of computer-aided tumor volume determination", *Journal of Surgical Research*, vol 59(6), pp. 621-626, 1995.

[6] J. K. Udupa and S. Samarasekera, "Fuzzy connnectedness and object definition: Theory, algorithms and applications in image segmentation", *Graphical Models and Image Processing*, vol 58(3), pp. 246-261, May 1996.

[7] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford Press, New York, NY, 1995.